

Лабораторная работа № 3 «Однофакторный дисперсионный анализ»

студента Яковлева Андрея группы Б20-504. Дата сдачи: 20.11.22

Ведущий преподаватель: Трофимов А.Г. оценка: _____ подпись: _____

Вариант № 12

Цель работы: изучение функций Statistics and Machine Learning Toolbox™ MATLAB / Python SciPy.stats для проведения однофакторного дисперсионного анализа (*One-Way ANOVA*).

1. Исходные данные

Характеристики наблюдаемых случайных величин:

СВ	Распределение	Параметры	Математическое ожидание, m_i	Дисперсия, σ_i^2	Объем выборки, n_i
X_1	$N(\mu, \sigma^2)$	$\mu = 5$ $\sigma^2 = 3$	5	3	200
X_2	$N(\mu, \sigma^2)$	$\mu = 5$ $\sigma^2 = 1$	5	1	250
X_3	$N(\mu, \sigma^2)$	$\mu = 5$ $\sigma^2 = 5$	5	5	200
X_4	$N(\mu, \sigma^2)$	$\mu = 5$ $\sigma^2 = 5$	5	5	200

Количество случайных величин $k = 4$

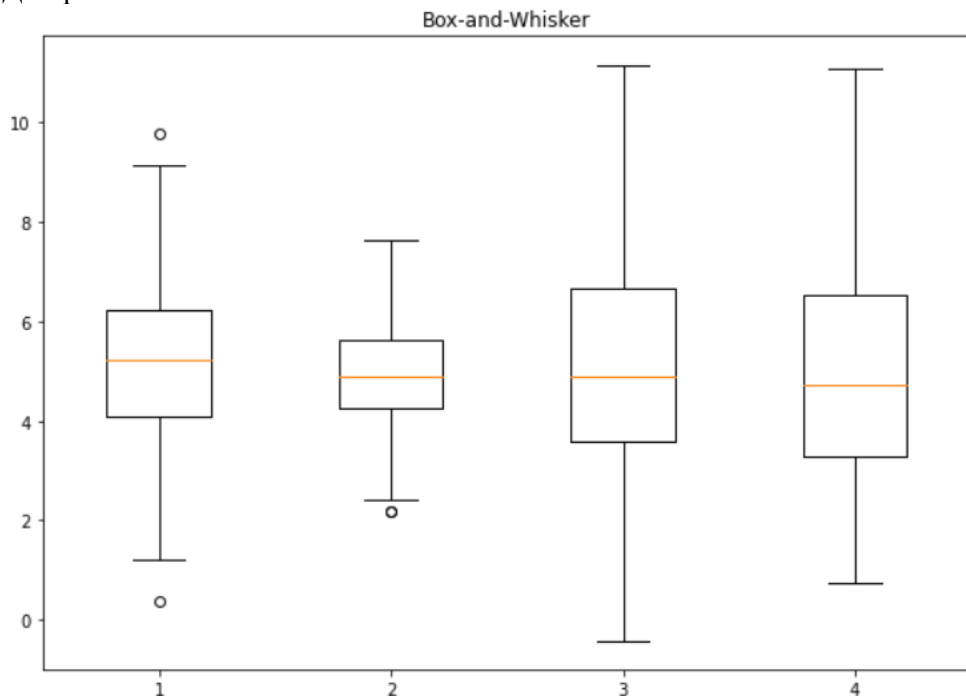
Примечание: для генерации случайных чисел использовать функции **rand**, **randn**, **chi2rnd** (scipy.stats: **uniform.rvs**, **norm.rvs**, **chi2.rvs**)

Выборочные характеристики:

СВ	Среднее, \bar{x}_i	Оценка дисперсии, s_i^2	Оценка с.к.о., s_i
X_1	5.162	2.926	1.711
X_2	4.879	0.982	0.991
X_3	5.064	4.795	2.190
X_4	4.967	4.419	2.102
<i>Pooled</i>	5.010	3.157	1.777

2. Визуальное представление выборок

Диаграммы *Box-and-Whisker*:



Примечание: для построения диаграмм использовать функции **boxplot**, **vartestn** (`matplotlib.pyplot.boxplot`)

3. Проверка условия применимости дисперсионного анализа

Статистическая гипотеза: $H_0 : \sigma_1^2 = \dots = \sigma_k^2$

Критерий Бартлетта:

Выборочное значение статистики критерия	<i>p-value</i>	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
152.435	$7.86e^{-33}$	H_0 отвергается	нет

Примечание: для проверки гипотезы использовать функцию **vartestn** (`scipy.stats.bartlett`)

4. Однофакторный дисперсионный анализ

Таблица дисперсионного анализа:

Источник вариации	Показатель вариации	Число степеней свободы	Несмещённая оценка
Группировочный признак	0.01179	$k - 1 = 3$	3.280
Остаточные признаки	3.145	$N - k = 846$	3.160
Все признаки	3.157	$N - 1 = 849$	3.161

Эмпирический коэффициент детерминации $\eta^2 = \underline{0.00367}$

Эмпирическое корреляционное отношение $\eta = \underline{0.06055}$

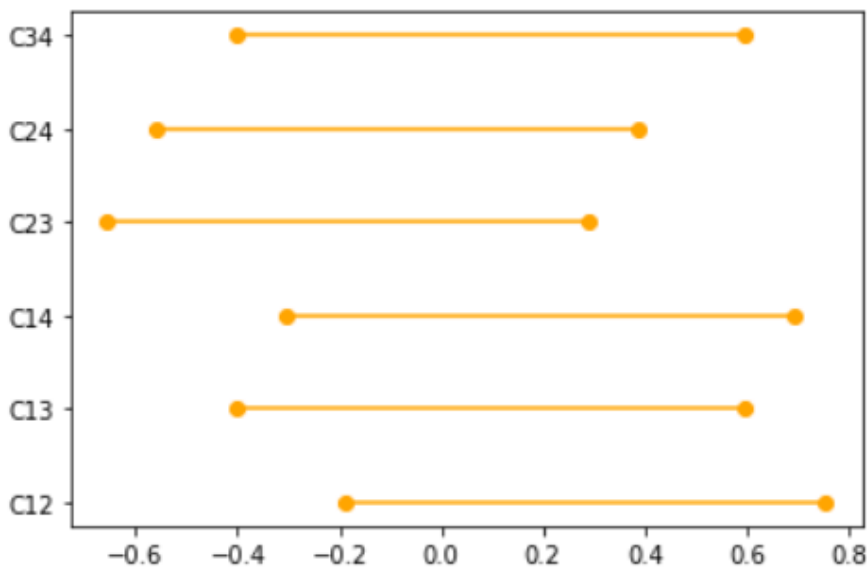
Статистическая гипотеза: $H_0 : m_1 = \dots = m_k$

Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.05$	Ошибка стат. решения
1.038	0.375	H_0 принимается	нет

Примечание: при расчетах использовать функцию **anova1** (**scipy.stats.f_oneway**)

5. Метод линейных контрастов

Доверительные интервалы для m_1, \dots, m_k :



Попарные сравнения m_i и m_j :

Гипотеза	Выборочное значение статистики критерия	p -value	Статистическое решение при $\alpha = 0.01$	Ошибка стат. решения
$H_0: m_1 = m_2$	2.186	0.0287	H_0 принимается	нет
$H_0: m_1 = m_3$	0.495	0.620	H_0 принимается	нет
$H_0: m_1 = m_4$	1.014	0.310	H_0 принимается	нет
$H_0: m_2 = m_3$	-1.188	0.234	H_0 принимается	нет
$H_0: m_2 = m_4$	-0.581	0.561	H_0 принимается	нет
$H_0: m_3 = m_4$	0.452	0.651	H_0 принимается	нет

Примечание: при расчетах использовать функцию **multcompare** (statsmodels.stats.multicomp.pairwise_tukeyhsd)